

Offprint from
**OXFORD STUDIES
IN ANCIENT
PHILOSOPHY**

EDITOR: BRAD INWOOD

SUPPLEMENTARY VOLUME

2012

VIRTUE AND HAPPINESS:
ESSAYS IN HONOUR
OF JULIA ANNAS

EDITED BY RACHANA KAMTEKAR

OXFORD
UNIVERSITY PRESS

2012

PSYCHOLOGICAL EUDAIMONISM AND INTERPRETATION IN GREEK ETHICS

MARK LeBAR AND NATHANIEL GOLDBERG

Do all men wish to do well? Or is this question one of the ridiculous ones I was afraid of just now? I suppose it is stupid even to raise such a question, since there could hardly be a man who would not wish to do well.

(PLATO, *Euthydemus* 278 E 3–5)¹

NO ONE has done more than Julia Annas to reintroduce ancient Greek eudaimonism into philosophical thought and discourse. Annas has (for good reason) focused on eudaimonism's *normative* work. The protreptic passage in *Euthydemus* which our epigraph begins, for example, is dedicated to establishing the conclusion that we should all seek 'good sense and wisdom' if we wish to live and do well.² What has received less attention is the kind of premiss to which Plato has helped himself here. He has Socrates make, rather casually, a bold and surprising *descriptive* claim, about the psychological—more specifically motivational—economies of human beings, asserting that within each such economy there is a powerful desire to live well. The argumentative context of these protreptic passages makes it clear that Plato thinks, not only that nobody lacks this desire, but also that this desire can reliably be

© Mark LeBar and Nathaniel Goldberg 2012

We are grateful for comments and discussion from participants in the 15th Annual Arizona Colloquium in Ancient Philosophy, including and especially Rachel Singpurwalla, our colloquium commentator, and Julia Annas, who first made us think about psychological eudaimonism in the context of ancient ethics. We are also grateful for comments and discussion from Robert Briscoe, Jeremy Morris, Alan Pichanick, James Petrik, Matthew Rellihan, Daniel Russell, Lad Sessions, Angela Smith, Florentien Verhage, and Tad Zawidzki. Finally, we would like to thank Washington and Lee University for awarding Goldberg a Lenfest Summer Grant to support this research.

¹ Trans. R. K. Sprague, in *Plato: Complete Works*, ed. J. M. Cooper (Indianapolis and Cambridge, 1997). All quotations from Plato come from this volume. See also *Euthyd.* 282 A.

² *Euthyd.* 281 B 8–282 A 1. 'Do well' translates *eu prattein*, which, like *eu zēn* (lit. 'live well'), is used as a synonym for *eudaimonia*, as the context makes clear.

counted on (when accompanied by correct beliefs about the role of morality or virtue in living well) to move people to seek ‘good sense and wisdom’—to be virtuous. Call the claim that this desire is found universally in human psychology, with this sort of dominating motivational force, ‘psychological eudaimonism’ (‘PE’).

PE appears not only in Plato but also in Aristotle and the Stoics. Nevertheless, we might wonder about it in two ways. First, we might wonder about its warrant. Though universal in scope, the claim concerns an apparently contingent feature of the motivational propensities of individual human organisms. What grounds could the ancients have for their confidence that there are no outliers?³ Second, we might wonder about PE’s relationship to *normative* eudaimonism. By ‘normative eudaimonism’ (‘NE’) we mean the normative or prescriptive claim that we have conclusive *reason* to act in ways that conduce to our own *eudaimonia*. NE is a claim about what it is to be rational.

We think both questions can be illuminated by considering the context of interpersonal interpretation in which ascriptions of mental states—including the dominant desire to live well—occur. Considering what is involved in seeing others as rational yields one essential clue as to why PE must be not only warranted but also true. Considering how the aspiration to live well contributes to our rationality yields another such clue. In fact we shall argue that PE is warranted by the conjunction of two ideas suggested by what is involved in understanding others as rational generally and ourselves as practically rational specifically. We take the former to be—interestingly and unexpectedly—informed by Donald Davidson’s account of how we interpret others. We take the latter to be the ancients’ account of NE. If we couple Davidson’s philosophical strictures on interpretation with the eudaimonist structure of practical rationality essential to ancient Greek ethics, then PE must be true.⁴

³ We assume that they do not merely misspeak in framing general claims as universal.

⁴ There are three reasons why bringing Davidson into contact with ancient philosophy is not as strange as it might seem. (1) Davidson’s undergraduate training was in classics, and he wrote his Ph.D. dissertation on *Philebus*, which he translated himself (Davidson, ‘Intellectual Autobiography of Donald Davidson’, in *The Philosophy of Donald Davidson* [Davidson], ed. L. E. Hahn (La Salle, Ill., 1999), 3–70 at 27). (2) Though Davidson published in philosophies of action, language, and mind, even here ancient ideas percolate through. As explained below, he published on *akrasia* (‘How is Weakness of the Will Possible?’, in *Essays on Actions and Events* [Essays],

The paper proceeds as follows. In Section 1 we provide a textual basis for ascribing PE to Plato, Aristotle, and the Stoics.⁵ In Section 2 we outline our two-part strategy for defending PE. The first part comes, in Section 3, from Davidson's account of interpretation. The second comes, in Section 4, from the ancients' account of NE. In Sections 5 and 6 we consider objections.

1. Psychological eudaimonism in ancient ethical theories

Where does PE appear in ancient Greek ethical theories? Our epigraph is the foundational premiss for Plato's argument in *Euthydemus* that we should seek to become wise. It is as clear as any formulation of PE we get in Plato. Similarly, in *Symposium* Plato treats the claim that happiness is universally desired as the terminus of a regress. He has Diotima contend that we want everything else for the sake of happiness, but there is 'no need to ask further, "What's the point of wanting happiness?"'.⁶ This point matters, as it indicates that, indeterminate as the end under discussion might be (she and Socrates have not yet agreed on what happiness consists in), it nevertheless plays a certain functional role in people's psychological economies. It is that *for the sake of which* they want other good things: it gives a 'point' to seeking them. And, as in *Euthydemus*, this desire is thought to be universal. She asks Socrates whether 'this desire for happiness, this kind of love' is 'common to all human beings', and whether 'everyone wants to have good things for ever and ever'. 'It is common to all', he agrees.⁶ In *Meno* the tenet that nobody wishes to be 'miserable and unhappy' grounds an argument for the value of virtue (78 A 3). And the vexed argument of

2nd edn. (New York and Oxford, 2001), 21–42). He elsewhere counselled that to find a speaker intelligible, we must consider her 'a believer of truths, and a lover of the good' ('Mental Events', *ibid.* 207–28 at 222); few statements sound more Platonic. Davidson even tries reading some of his views on interpretation into Plato (*Truth, Language, and History* (Oxford, 2005), essays 15–18). (3) Davidson and the ancients both share a methodological commitment to the primacy of the first-personal agential perspective. We say more about this in sect. 4.

⁵ For our purposes these three schools comprise 'ancient Greek ethics'. We recognize, of course, that there were others. But only these three focus on *agency* in the good life of the sort necessary to secure our argument.

⁶ *Sym.* 205 A 4–10, trans. Nehamas and Woodruff. See also 205 D 2–4: 'every desire for good things or for happiness is "the supreme and treacherous love" in everyone'. In both passages 'happiness' translates *eudaimonia*.

Philebus is based on the belief that ‘everything that has any notion of [the good] hunts for it and desires to get hold of it and secure it for its very own, caring nothing for anything else except for what is connected with the acquisition of some good’ (*Phileb.* 20 D 6–8, trans. Frede).⁷

Aristotle’s ethical theorizing likewise depends on the claim that is formulated as an explicit commitment in the *Nicomachean Ethics*: ‘[H]appiness is among the things that are prized and complete. . . . [I]t is for the sake of this that we all do everything else.’⁸ That too is about as explicit a formulation of PE as one might desire. Although Aristotle does not make the idea explicit in his discussion of voluntary and involuntary actions, PE is also at work in his exploration of the degree of voluntariness in actions adverse to one’s own interests.⁹ Additionally, PE is implicit in the argument Aristotle offers in the *Eudemian Ethics* for identifying the normative aim of life; this, he maintains, is ‘most difficult’ when it comes to what is desirable, what would satisfy our desire (*EE* 1. 5, 1215^b15–19, trans. Solomon). Aristotle clearly supposes that the dominant desire to live well is something common to, and descriptive of, ‘the multitude’ (*EE* 1. 3, 1214^b34).

Finally, PE is endorsed by the Stoics as well. Epictetus relies on it to argue that one ‘cannot think a thing advantageous to him, and not choose it’.¹⁰ PE is also key to the Stoic doctrine of ‘personal *oikeiōsis*’. The core of this doctrine is that each of us is born with a drive to seek what is good for us. The objects of this drive begin, as we are children, with food and warmth; we start to use reason to attain our objects, and eventually learn to treasure reason for its own sake. On the Stoic view the good life *just is* the life according to nature, so—since the ‘impulse’ given by nature occurs in each of us—a form of PE lies in the background of the entire doctrine.¹¹

⁷ Though this claim is framed as one about the desire for ‘the good’ (*to agathon*), the context makes it clear that the good in question is a property of one’s life. The good life is the quarry of the dialectic.

⁸ *NE* 1. 12, 1102^a1–3, trans. D. O. Ross, rev. J. O. Urmson, in *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. J. Barnes, 2 vols. (Princeton, 1984). All quotations from Aristotle come from this volume.

⁹ e.g. throwing one’s valuables into the sea to avoid drowning in a storm (*NE* 3. 1, 1110^a5).

¹⁰ Epict. *Diss.* 1. 28. 6, in *The Discourses of Epictetus. The Handbook. Fragments*, ed. C. Gill, trans. R. Hard (London, 1995).

¹¹ For ‘personal *oikeiōsis*’ see e.g. Cic. *Fin.* 3. 33; 3. 59; 4. 16; 4. 34; 5. 34. On the good life as life according to nature see e.g. *Fin.* 3. 26; 3. 31.

This gets us to our central question, as to the warrant for PE itself in these theories. PE is a descriptive claim about the psychology of individual human beings. But there seems to have been no systematic investigation of any empirical facts to sustain it, and though its truistic nature goes some of the way to do so, it certainly doesn't rule out outliers as the Greeks seem to suppose it does. Can PE be warranted as a genuinely universal claim? Can it also be true? Next we lay out our strategy for arguing that it is both.

2. Defending psychological eudaimonism

We propose that PE's warrant arises from a *transcendental* claim about the possibility of engaging in argument about how to live and act. We have in mind a claim about the possibility of 'interpretation'—of ascribing intentional mental states to those with whom we engage in discourse. To see this let us begin by making PE precise:

(PE) Each person has the desire to live well as a dominant element in his or her psychological economy.

On our view, the ancients (and for that matter we ourselves) are justified in asserting PE in virtue of the truth of a closely related *transcendental* (T) presupposition of PE:

(TPE) If we are to ascribe rationality to someone, we must also ascribe to him or her the desire to live well as a dominant element in his or her psychological economy.

We must consider how TPE can supply warrant to PE. We do so after two clarificatory comments about it.

First, by 'rationality' we mean both theoretical and practical. At stake is the possibility of engagement in argumentation about action. The interpretation involved is as of another person capable of assessing reasons for and against possible courses of conduct. Second, by speaking of a desire to live well as a 'dominant element' in the psychological economy of a person (which we shall hereafter speak of as having the 'dominant desire'), we mean two things specifically. First, it must be motivationally powerful enough to be called upon to provide sufficient motivation to overwhelm competing mo-

tivations. Second, it must be stable and enduring enough to be called upon over time and throughout the process of habituation.¹²

Now TPE, as a transcendental claim as to the very possibility of interpreting another as rational, turns out to be both necessary and normative. It is necessary because (we shall argue) it applies in virtue of our concept of rationality, correctly understood. It applies in virtue of what rationality turns out (under reflective scrutiny) to be. It is normative because it prescribes the conditions under which (we shall argue) we may, and may not, appropriately ascribe rationality to others.

PE, as a descriptive claim, is not the same as TPE's consequent. How can TPE warrant PE? TPE's conclusion is a *prescription*: a claim about what we *must* do if we are to interpret others as rational. TPE, if true, *licenses* an ascription of the dominant desire to live well to anyone being interpreted as rational. And interpretation as rational is necessary for anyone to whom argument is directed; we could not understand ourselves as *engaging in argument* except that we see those we engage as susceptible to the normative constraints of rationality. This was true of Socrates and the ancient eudaimonists, and it is true of us. The truth of TPE ensures that in any context in which PE is deployed, its deployment will be warranted in virtue of rationality of those to whom the argument is directed.

Would TPE secure the *truth* of PE? It must. PE makes a universal claim, about 'each' (and every) person. We can understand TPE as making explicit the domain of that claim, such that it is indeed universally true. It specifies that the domain of 'persons' for the purposes of PE is all and only those beings we interpret as rational. If we take the antecedent of TPE to be 'fixing the domain' of persons in this way, then its truth guarantees the truth of PE also.

The question that will be before us for the rest of this paper is *whether* TPE is true. Our argument that it is has two premisses. Consider how things look to Socrates. First, to understand others as rational generally, Socrates must interpret them as similar to himself in his own rationality. He must see them as having a psychological economy with beliefs, desires, and evaluative attitudes in

¹² When we speak of a 'desire', we mean only an attitude characterized by motivational valence (attraction or aversion). We remain neutral on whether we desire only what we judge to be the good. See R. Kamtekar, 'Plato on the Attribution of Conative Attitudes', *Archiv für Geschichte der Philosophie*, 88 (2006), 127–62, for an argument that in attributing such conative attitudes to others we must see them as aiming at the good.

the relevant respects like his own—as having the intentional states that he would have were he in their place. So must we. This is tantamount to Davidson’s argument for the charity constraint on interpretation. Call this premiss ‘Davidson’s Principle of Interpretation’ (‘DPI’):

(DPI) We can interpret others as rational only by ascribing to them intentional states that we would have were we in their place.

DPI implies that we can fail to see others’ practical agency as like our own only on pain of ceasing to see them as rational. *Second*, Socrates could not understand himself as (specifically) practically rational without the dominant desire to live well. Neither can we. This comes from what we call the ‘Eudaimonist Principle of Rationality’ (‘EPR’):

(EPR) The dominant desire to live well is essential to understanding ourselves as practically rational.

By ‘essential’ we mean that we could neither be practically rational nor understand ourselves to be practically rational without it; this is something we discover when we reflect on what living well consists in, and in particular consider the role of virtue and successful practical rationality (practical wisdom) in living well. Because the dominant desire to live well is essential to *understanding* ourselves as practically rational, it is essential to our intentional states overall.

Hence DPI tells Socrates that to understand others as rational he must consider what his own intentional states would be in the relevant situation and project them outward. EPR tells Socrates that his own intentional states must contain the dominant desire to live well. Now if Socrates can interpret others as rational only by ascribing to them intentional states that he would have were he in their shoes, *and* the dominant desire to live well is essential to his understanding of himself as practically rational and so essential to his own intentional states, then Socrates must ascribe this desire to them as well on pain of ceasing to see them as rational.

We may think of DPI as providing the form that interpretation of others as rational must take (roughly, our own), while EPR provides the content of practical rationality in particular (crucially, including the dominant desire to live well). DPI by itself does not establish that speakers share any specific desire, let alone the dominant de-

sire to live well. Davidson himself acknowledges as much.¹³ Nor do the ancient eudaimonists claim that it is conceptually impossible for rational beings *in abstracto* to lack the dominant desire to live well. Their commitment to EPR is a commitment to how *we* are to understand *ourselves*. Though we can understand others on the model of ourselves, nothing about EPR sanctions the conclusion that it is impossible to do otherwise. That is why DPI and EPR are both needed to ensure the truth of TPE. And if TPE is true, then the truth of PE follows.

Now we do not claim that Socrates (or anyone else) consciously applies DPI or EPR to others when attempting to understand them. Our point instead is that we as theorists necessarily could recover from any agent's acts both DPI and EPR as principles constitutive of how that agent understands the acts of others. The theoretical ascription of DPI and EPR to an agent, even if the agent is not herself consciously guided by them, is the only means by which we could understand how one agent can in principle make sense of another. In the next two sections we rehearse DPI, EPR, and their conclusion.

3. Donald Davidson on interpretation

Davidson's canonical work on interpretation begins, 'What is it for words to mean what they do?'¹⁴ He answers in two stages.¹⁵ First, Davidson argues that words mean what they do in virtue of their role in a speaker's sentences, and those sentences mean what they do in virtue of the meaning that would be assigned to them by a radical interpreter. The radical interpreter, in turn, assumes nothing

¹³ See Davidson, 'The Interpersonal Comparison of Values', in *Problems of Rationality [Problems]* (Oxford, 2004), 59–74 at 72. His expositors agree. See B. Martin, 'The Moral Atmosphere: Language and Value in Davidson', *Southwest Philosophy Review*, 6 (1990), 89–97, and 'Interpretation and Responsibility: Excavating Davidson's Ethical Theory', in Hahn (ed.), *Davidson*, 345–56; C. Rovane, 'On the Very Idea of an Ethical Scheme', *Iyyun*, 54 (2005), 109–28; and D. Caldwell, 'The Shared-Value Thesis', *Southwest Philosophy Review*, 21 (2006), 61–8.

¹⁴ D. Davidson, *Inquiries into Truth and Interpretation [Inquiries]* (Oxford, 2001), xiii.

¹⁵ The first stage is constituted especially by *Inquiries*, essays 1–5, 9–12, and *Subjective, Intersubjective, Objective [Subjective]* (Oxford, 2002), essay 10; the second, by *Problems*, essays 1–4, and *Truth and Predication* (Harvard, 2005), chs. 1–3. Our account of these is necessarily limited. For other accounts specifically of the first, see N. Goldberg, 'Tension within Triangulation', *Southern Journal of Philosophy*, 46 (2008), 367–83, and 'Triangulation, Untranslatability, and Reconciliation', *Philosophia*, 37 (2009), 261–80.

about what a speaker's words mean. Instead, she determines what they mean by constructing a charitable, Tarski-style truth theory for the speaker's language. In virtue of being a Tarski-style truth theory, the theory systematically correlates a speaker's utterances with conditions under which each is true.¹⁶ In virtue of being charitable, the theory requires that the interpreter make these truth-value assignments given what she herself would believe were she in the speaker's place.

Because the charity condition is central to our argument, it requires elaboration. According to Davidson, what a speaker utters (at least in basic cases)¹⁷ depends on both what she takes her words to mean and what she believes about her environment. The only way in which a radical interpreter can make the requisite truth-value assignments in these basic cases that will allow her to solve for meaning is if she holds the content of the speaker's beliefs constant, as far as possible. Davidson's proposal is that the interpreter do so by taking those beliefs, as far as possible, to be true. Now, that is true *as taken by the interpreter*. According to Davidson, interpretation can proceed only if the interpreter ascribes to the speaker as much as she can believe that she *herself* would have were she in the speaker's spot.

Suppose the radical interpreter observes a speaker utter 'Lah' when and only when it is snowing. Unless the interpreter ascribes to the speaker the *belief* that it is snowing (which the interpreter herself would have were she in the speaker's place), she has no reason to interpret 'Lah' to *mean* that it is snowing. Without assuming that the speaker believes things that the interpreter herself would, the interpreter has no reason not to take the speaker to mean that it is raining, that it is sunny, that boulders are falling from the sky, or anything else. Further, Davidson stresses, beliefs can themselves be ascribed only against a network of related beliefs. And these related beliefs, like the initial ones, are drawn from the interpreter's own potential stock. Unless the interpreter *also* ascribes to the speaker beliefs like these—that snow falls from the sky and can accumulate on the ground, that snow is distinct from dirt, etc. (which the interpreter would also believe in the speaker's situation)—she is not entitled to interpret 'Lah' to mean it is snowing, either. In fact,

¹⁶ See Davidson, 'Moods and Performance', in *Inquiries*, 109–22, for utterances that are not statements.

¹⁷ Such cases, for Davidson, are presumably determined via trial and error.

without believing these (and other) things about snow, the speaker cannot be taken to believe *anything* about snow. 'The reason for this', Davidson explains,

is that a belief is identified by its location in a pattern of beliefs; it is this pattern that determines the subject matter of the belief, what the belief is about. Before some object in, or aspect of, the world can become part of the subject matter of a belief (true or false) there must be endless true beliefs about the subject matter. False beliefs tend to undermine the identification of the subject matter; to undermine, therefore, the validity of descriptions of the belief as being about the subject.¹⁸

Truth or falsity, that is, by the interpreter's lights.

Now Davidson counsels two caveats. First, when the radical interpreter interprets a speaker's utterances, she may occasionally have to ascribe to the speaker beliefs that the interpreter would regard as false. Suppose the interpreter produces a truth theory according to which, when and only when it is snowing, the speaker utters something that means that it is raining. The charity condition would require the interpreter to revisit her initial belief ascription to avoid taking the speaker to say things that, when and only when it is raining, *mean* that it is snowing. (Perhaps the interpreter would now ascribe the belief that snow is a kind of ubiquitous moisture that only sometimes forms flakes.) The best way *overall* for the interpreter to ascribe her own beliefs to the speaker is occasionally to ascribe beliefs that the interpreter herself would lack were situations switched.

Davidson's second caveat is that interpretation is indeterminate. There will always be more than one way to interpret a speaker's utterances, because there will always be more than one belief to ascribe to any speaker given any situation. We saw that above with ascribing beliefs about snow and rain. None the less, indeterminacy or not, interpreters aim for a theory that *best* reconciles charity with the formal conditions of a Tarski-style truth theory. 'Nothing more is possible, and nothing more is needed'¹⁹ to determine what a speaker's words mean.

Desires come in the second stage of Davidson's work on interpretation. The radical interpreter this time begins by isolating utterances that the speaker *prefers* to be true, for Davidson argues that

¹⁸ Davidson, 'Thought and Talk', in *Inquiries*, 155–70 at 168.

¹⁹ Davidson, 'On the Very Idea of a Conceptual Scheme', in *Inquiries*, 183–98 at 197.

a speaker's preferences are determined by her beliefs and desires together. The interpreter now begins by charitably ascribing beliefs and desires to determine what those words mean. She takes her speaker to believe what is true (by her lights) *and* desire what is good (by her lights) too.

Suppose the radical interpreter observes a speaker utter 'Jitwat ugum' and point at a bowl of sugar when and only when presented with a bowl of sugar and one of salt. Because, for Davidson, preferences resolve into beliefs and desires, the interpreter must (as far as possible) hold *both* constant to solve for meaning. Davidson's proposal is to do so by ascribing to the speaker (as far as possible) beliefs and desires *together* that the interpreter would have were she in the speaker's place. Unless the interpreter ascribes to the speaker (a) the *belief* that she is being presented with a bowl of sugar and one of salt, and that she is pointing at the former, and (b) the *desire* to have sugar over salt, she is not entitled to interpret 'Jitwat ugum' to *mean* that I (the speaker) prefer(s) sugar over salt, or something like it. Moreover, unless the interpreter also ascribes to the speaker desires like these—to have sugar over nothing, to have others pay attention to one's gesticulations, etc. (which the interpreter would desire also in the speaker's place)—she is unjustified in interpreting 'Jitwat ugum' to mean that I (the speaker) prefer(s) sugar over salt, or something like it, either. Desires can be ascribed only against a network of desires. In fact Davidson maintains that beliefs and desires can be ascribed only collectively against a network composed of both. And ascriptions of both would, again, derive from the interpreter's self-reflection. Finally, these assumptions on hand, the interpreter can take 'Jitwat ugum' to mean that I (the speaker) prefer(s) sugar over salt. And that is enough, according to Davidson, for it to have that meaning.²⁰

Now all this is *radical* interpretation. The radical interpreter, however, is meant to be an idealized version of *us*. In practice we all assume much about the meaning of a speaker's words; none of us constructs a charitable, Tarski-style truth theory for a speaker's language. But in principle, according to Davidson, we all could. Reflecting on the procedure of the radical interpreter is meant to reveal

²⁰ See N. Goldberg, 'Swampman, Response-Dependence, and Meaning', in G. Preyer (ed.), *Davidson's Philosophy: Truth, Meaning, and the Mental* (Oxford, forthcoming), for why this commits Davidson to a response-dependence account of meaning.

how belief, desire, meaning, and truth can generally be teased out of speaker preferences and interpreter assumptions.

To see this, consider Andrea, a hard-working lawyer, devoted wife, and dedicated mother. We observe Andrea saying things like:

- (α) 'I prefer getting to the office tomorrow earlier rather than later.'
- (β) 'I prefer that my husband do the grocery shopping.'
- (γ) 'I prefer red over white wine with dinner.'

Of course, she says things that are not statements of her preferences; and some things that she says we don't hear. But we do hear these. Since we are not radical interpreters, we assume that Andrea's words mean the same as ours. But (and this is Davidson's point) if Andrea turned out to believe and desire things very different from what we would in her place, we would have no warrant to assume that we knew *anything* about what her words mean. What entitles *us* to interpret (α) to mean that I (she) prefer(s) getting to the office tomorrow earlier rather than later is that we take Andrea to believe and desire generally what we would in her place, including, perhaps, the belief that it makes more sense to get there earlier and the desire to do what makes more sense. We can interpret (β) to mean that I (she) prefer(s) that my (her) husband does the grocery shopping only if we again ascribe to her beliefs and desires that we ourselves would have were we in her spot, perhaps this time including the belief that it would be better for him to do the shopping and the desire to do what is better. We can interpret (γ) to mean that I (she) prefer(s) red over white wine with dinner only if we ascribe to her what we would believe, maybe that either kind of wine is optional, and what we would desire, maybe to enjoy what I (she) drink(s). Hence we are ascribing to Andrea the desires:

- (1) to do what makes more sense;
- (2) to do what is better;
- (3) to enjoy what I drink.

Admittedly these are not the only desires that we could ascribe to Andrea, on Davidson's view. Even non-radical interpretation retains an element of indeterminacy: we might ascribe to her different beliefs also and interpret her words to have different meaning. None the less, the differences are not limitless. Given our observations of Andrea, and our general ability to make sense enough of her

behaviour as our theory of what her words mean evolves, we would not ascribe to her these desires:

- (1') not to do what makes more sense;
- (2') not do what is better;
- (3') not to enjoy what I drink.

As interpreters we have no starting-point other than ourselves. Consequently we have no better indicator of what someone else would believe or desire in various situations than what *we* would. And, given Andrea's behaviour, (1')–(3') would not match our conative state were we in her situation. We would therefore have no reason to ascribe (1')–(3') to her.

Though we reserve judgement on much of Davidson's proposal, we agree that the only way in which an interpreter can assume (as we do) *or* determine (as the radical interpreter does) what a speaker's words mean is by ascribing to the speaker the intentional states that we would have were we in her situation—readjusting as necessary, with an eye towards taking her overall mental stock to match what would be ours. Hence regardless of how much or little an interpreter assumes about what a speaker's words mean, the interpreter has no choice but to assume a great deal about what the speaker believes and desires. In fact Davidson concludes:

A creature that cannot in principle be understood in terms of our *own* beliefs, values [or desires], and mode of communication is not a creature that may have thoughts radically different from our own; it is a creature without what we mean by 'thoughts'.²¹

Therefore, we can interpret others as rational only by ascribing to them intentional states that we would have were we in their place. This is DPI.

4. The dominant desire to live well and normative eudaimonism

Suppose Davidson is correct and DPI is true. Why should TPE follow? Why must Socrates (say) have the particular desire to live well, let alone have it be *dominant*? The key is the role of the dominant

²¹ Davidson, 'Expressing Evaluations', in *Problems*, 19–38 at 37 (emphasis added). For Davidson's use of this reasoning against conceptual relativity, see N. Goldberg, 'E pluribus unum: Arguments against Conceptual Schemes and Empirical Content', *Southern Journal of Philosophy*, 42 (2004), 411–38.

desire to live well in Socrates' apprehension of himself as practically rational. The key claim is what we have called the 'Eudaimonist Principle of Rationality':

(EPR) The dominant desire to live well is essential to understanding ourselves as practically rational.

Here we adduce two arguments for EPR. The first is found in Aristotle's work on behalf of NE; the second in Plato's.

Aristotle's case in the opening chapters of the *Nicomachean Ethics* is that an ultimate end—one which lends reason-giving force to all other ends and means to them, but which is given no reason-giving force by any further end—is necessary to make rational the economy of end-seeking attitudes and conduct characteristic of the only practically rational beings we know, viz. ourselves. Seeing ourselves as having such an end is essential to understanding *ourselves* as practically rational. Moreover, Aristotle (like Socrates, Plato, and the Stoics) maintains that the only end suitable to that job is that of living well (*NE* 1. 12, 1102^a3). That is NE. Put differently, we have many final ends (ends we seek for their own sakes, not merely for the sake of something further), but all our final ends *except* the end of living well we have partially for the sake of a further end, viz. the end of living well. The ancients believe that an intelligible framework of human practical rationality must take the form of a hierarchy of ends, and the end that justifies and regulates the normative force of these others is the end of living well.

Why must there be a *single* end terminating all these chains of means–end justification? The answer depends both on what practical rationality looks like for the ancients, and on our understanding of its being fixed (as interpretation is) ineluctably in the first-person perspective of practical agency, rather than from the third-personal perspective of observation or explanation.

Consider Annas's response to the question. Annas claims that, in supposing that we must converge on a single ultimate end, Aristotle is being faithful to psychological fact, rather than making a point about rationality. She observes that Aristotle even acknowledges the existence of people who think that happiness is a 'compound' of several such ends, but then remarks:

These people, however, are presumably unreflective people who have not yet thought through the implications of reflection on one's life as a whole . . .

it seems clear that they take it that one cannot consistently stop at two or more ends if one is thinking of one's life as a whole. Felt discomfort with the idea of two or more uncoordinated ends will pressure the agent to continue towards a single final end.²²

Annas is right about 'felt discomfort', in particular in thinking about one's whole life, but we experience that discomfort precisely *because* we are practically rational, and conditions in which there is no ultimate end to adjudicate justificatory conflicts are conditions in which practical rationality is impossible: our decision in the face of such conflicts must be arbitrary.²³ To have an end is, *inter alia*, to accept normative constraints on the conditions under which one is willing to pursue alternatives. Ends impose normative constraints on those who can recognize them. Because we can make sense of our lives *as* the lives of practically rational creatures *only by* appreciating the role and force of ends from the first-personal perspective of agents whose ends they are, Aristotle's claims about the work of an ultimate end in our practical economies are rooted in our appreciation of the normative dimensions of those economies.

Of course, his proposal is not that people operate with this end-seeking structure consciously present to mind, let alone with what Sarah Broadie calls 'The Grand End' in view.²⁴ The structure is normative, becoming explicit only in rational reflection. As Annas reminds us, for the Greek theorists the aim of 'ethical reflection' was the concern to make one's life as a whole a good one. Such reflection does two things. First, in resolving the normative tensions between our goals and aspirations, it renders determinate the nature of the end of living well. If the ancients are right, reflection reveals that virtue is at least necessary, if not also sufficient, for a good life. Lives oriented towards pleasure, honour, or other candidate ultimate ends do not withstand scrutiny. Second, the process of reflection makes manifest the fundamental justificatory work of the end of living well in the life of the practically rational human being. No other structure of ends can be stable against the sorts of demand to which human practical rationality is subject.

Nevertheless, so much gets us only to the *end*, not the *desire*, of

²² J. Annas, *The Morality of Happiness* (Oxford, 1993), 33. The passage in which Aristotle remarks on those accepting a 'compound' is at *EE* 1. 1, 1214^b4.

²³ For a defence of this claim see M. LeBar, *The Value of Living Well* [*Living Well*] (Oxford, forthcoming), chs. 1–2.

²⁴ S. Broadie, *Ethics with Aristotle* (Oxford, 1991), ch. 4.iv.

living well, as essential to practical rationality. These do not, however, readily divorce—on the ancients' views or on ours. For the ancients, the dominant desire to live well plays an essential role in the reflection that vindicates NE. First, its strength secures the end of living well against countervailing motivations. Given the motivation to act viciously in some way, and having grasped that vicious conduct is in conflict with the end of living well because of the necessity of virtue for living well, one might consider the option of giving up the end of living well. But given the contribution of virtue to living well, the cost of subordinating the dominant desire to live well to a desire to act viciously is a disavowal of the structure of ends and means that shapes one's reasons to act. The cost of doing so is therefore to cease being practically wise.

Second, and more importantly, the desire is essential to the end because of the relationship between desires and ends on the ancient ethical theories. Desires do not themselves establish ends, yet ends involve normative constraints and not simply conative states. A Kantian take on that point would be to say that desires require being taken up or 'incorporated' through some process of rational endorsement,²⁵ but this isn't quite adequate as a characterization of the ancient approach. For such endorsement is only one of the two ways in which desires and ends are connected. Desires do not simply come unbidden. *We* bid them in virtue of the kinds of people we make ourselves, through moral education and then habituation. We cultivate the desires we endorse, and those we do not endorse we attempt to extirpate by the same means.²⁶ The natures we cultivate are themselves the *sources* of desires we judge worth having.

From this perspective, the dominant desire to live well is one that ethical reflection sustains as one we have good reason to cultivate, suitable as an end. Because we desire it, we are motivated to pursue it even when we would prefer not to. Because we endorse it, it

²⁵ Kant says: 'Freedom of the power of choice has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim*' (Kant, *Religion within the Boundaries of Mere Reason*, 6. 24, trans. A. W. Wood and G. Di Giovanni, in *Religion and Rational Theology* (Cambridge, 1996)).

²⁶ A clear case of this view is contained in Socrates' argument against Callicles in Plato's *Gorgias*. Socrates exerts considerable energy to drive Callicles, who is defending the view that the best life is one of pleasure derived from satisfying desires, to agree that there are both good desires and bad (499 ff.). Once that is granted, the knowledge of which is which—and thus which we ought to seek to satisfy and which not—takes normative priority.

acquires not just conative but also normative force as an end that guides and constrains our actions. In so acting we develop natures that foster the perseverance of the desire. Our endorsement of the dominant desire to live well gives us reason to cultivate it in life and action, in part by gratifying it. The cost of lacking the dominant desire to live well, conversely, would be tension between our motivational propensities and the end we see as that for the sake of which we do all that we do. Such tension would undermine a disposition to regulate one's ends in a way that is stable under ethical reflection. It would be a threat to a stable disposition to comply with the demands of practical rationality itself.

A second argument for EPR comes from Plato's work, and this argument too (found also in Aristotle) binds the end and the desire in the practically rational agent and hence conjoins EPR and NE. The Greek eudaimonists argue for the end of living well as our ultimate end by positing *formal conditions* on the ultimate end that gives focus and structure to practical agency. The dominant desire to live well is crucial for these conditions. The work it does is most salient in the condition of *sufficiency* (*hikanon* in Plato, or *autarkes* ('self-sufficiency') in Aristotle). Our ultimate end, or good, must be plausible as an object of desire, indeed of a sort of comprehensive desire that dominates one's very life and agency. This line of thought appears in Plato's *Philebus*, where Socrates says:

Everything that has any notion of it hunts for [the good] and desires to get hold of it and secure it for its very own, caring nothing for anything else except for what is connected with the acquisition of some good. (*Phileb.* 20 D 6–8)

Here the dominant desire to live well is not framed as clearly as in *Euthydemus*, but the point is the same, for it is *lives* to which the test of sufficiency is applied. Socrates' interlocutor Protarchus agrees with Socrates that lives of pure pleasure and pure knowledge both fail the test, in that neither is a fit object of desire for human beings. Only a 'mixture' of the two can be endorsed, precisely because the alternatives do not satisfy our desire to live well. Self-sufficiency as a condition on our ultimate end does just the same work in Aristotle's argument for NE in *Nicomachean Ethics* 1. 7. These ancient arguments *bind together* the desire and the end. Responsiveness to the dominant desire to live well plays a crucial role as a normative constraint on what can count as our ultimate end. The dominant

desire to live well does important work, not only in the moral psychology of the virtuous agent, but also in the kind of philosophical rationale the ancients offered for their conceptions of the good life for the practically rational agents we are.

One might wonder at this point why DPI is needed to reach TPE, if EPR by itself entails a rational requirement to desire to live well. EPR is ultimately a principle about how we understand ourselves, from the first-personal perspective of practical agency. Though we can understand others as practically rational by taking them to desire to live well, nothing about EPR requires that we do so. Because EPR tells us that the dominant desire to live well is essential to understanding ourselves as practically rational, it follows that the same must be true if we are to understand others as practically rational on the *model* of ourselves. DPI supports the crucial point that that model is the only one available to us *qua* interpreters.

As illustration, reconsider the desires that we ascribed to Andrea:

- (1) to do what makes more sense;
- (2) to do what is better;
- (3) to enjoy what I drink.

Recall that we ascribed (1)–(3) instead of (1′)–(3′):

- (1′) not to do what makes more sense;
- (2′) not to do what is better;
- (3′) not to enjoy what I drink

because DPI required that we do so. Now suppose that *we* were Andrea, and *we* were asked why *we* uttered these:

- (a) ‘I prefer getting to the office tomorrow earlier rather than later.’
- (β) ‘I prefer that my husband do the grocery shopping.’
- (γ) ‘I prefer red over white wine with dinner.’

We would explain our goals or ends and, if pressed, arrive at a reflective chain of deeper explanatory and justificatory ends. We want to do what makes more sense when possible because we want to see what we do as meaningful, perhaps. We want to do what is better because we have the felt need to do what we think is right more often than not. We want to drink the red wine because we like red. Ancient eudaimonism is founded on the claim that once we have

begun this chain of reflection, it can terminate only with the aim of living well. That aim would give us reason to do all we do, including endorsing and acting upon desires like (1)–(3) rather than (1′)–(3′). That model of practical rationality is the only one available to us in interpreting Andrea. We can make sense of Andrea's desires as the desires of a rational person only by seeing them as in a framework in which the aim, and consequently the desire, to live well 'stands behind' such subordinate desires. Only the conjunction of DPI and EPR entitles us to interpret Andrea as having this desire. Only their conjunction allows us to interpret her words and deeds in a way that lets us make sense of her as practically rational.²⁷

Note that on our account of EPR practical rationality and our self-understanding as practically rational converge. The picture of practical rationality at work in the ancient eudaimonists is grounded in the first-personal perspective of agency, in which the experience of seeking ends, and responding to the experience of their normative constraints, is a familiar and irreducible part of practical agency. Davidson grounds his own account of interpretation in just such a first-personal perspective, as evidenced by DPI. The interpreter has no choice but to use her perspective as a starting-point to make sense of others. That is why the two approaches are so congenial.

In fact, if the ancients' approach is correct, we have reason to favour Davidson's privileging of the first-person agential (for him, interpretative) perspective over any third-personal scientific perspective that more empirically minded philosophers might prefer. When those such as Jerry Fodor think that identifying intentional states in human beings will ultimately turn out to be as objective, and so as amenable to third-person analysis, as identifying calcium (say) in certain minerals, they fail to take seriously enough the necessarily normative nature of the intentional. Even if intentional states in brains are as 'Real'—to use Fodor's emphatic term²⁸—as

²⁷ This is so even granted Davidson's indeterminacy doctrine. While we would not ascribe to Andrea (3′), because given the circumstances we would not desire that ourselves, we *might* ascribe to her this desire:

(3″) to drink what is allegedly heart-healthier.

Suppose that we did. Through the same reflective chain of deeper explanatory and justificatory ends we would desire this too because we desire to live well. Indeterminate or not, our ascription of *rationality* requires our ascription of the dominant desire to live well.

²⁸ J. A. Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass., 1989).

calcium is in certain minerals, their identification cannot help but appeal to the first-person perspective of what the states *should* be, given the interpretative circumstances. Normativity is part of our notions of beliefs and desires: it finds its home in each of us *qua* agents. That is why Davidson's first-person approach concerning meaning and mind is so congenial when reflecting on questions the ancients found fundamental, such as 'What kind of life is best?'. It is no coincidence that DPI and EPR can function in a single argument.

Still, we may doubt the truth of TPE, their conclusion. *Must* we think that all people, at least in so far as we can recognize them as practically rational, have the dominant desire to live well? We take up this in our final two sections.

5. Objections: Kant

One challenge to TPE comes from Kant. On his view, practical rationality is characterized by autonomy, in governing ourselves by self-legislated principles, rather than those imposed upon us by our heteronomous natures. Kant thinks that this requires that we exercise our wills in accordance with principles whose content is fixed formally, i.e. by the very idea of a principle drawing on no empirical incentives for its content. This is his Categorical Imperative. Its formal nature rules out any contribution from considerations relevant to creatures living contingent, organic, biological lives as we do, and thus offers the prospect of a model of practical rationality that can simply reject EPR and with it TPE.

Is this a reason for us to resist our argument for TPE? There are two reasons why it is not. First, since Hegel, there has been concern with this purely formal conception of practical rationality. If Kant cannot pull the substantive rabbit out of the formal hat, then this is no real alternative. Whether he can or not is a controversial question, which we cannot even begin to address here. Our point instead is that Kant's conception has at least as much doubt attending to it as either DPI or EPR.

Second, our argument for TPE may survive even if Kant is right: he may even count as a supporter. Kant holds that the 'principle of self-love' is present in not just all human beings but all 'dependent

rational beings'.²⁹ This principle is the imperative that they seek their own 'happiness'. Kant's justification for this claim is mysterious, by his lights, and indeed it is not easy to see what could account for its status as the 'imperative' he takes it to be.³⁰ Christine Korsgaard remarks that, in its nature as 'substantial' and not merely formal, it is a striking departure from the other imperatives Kant frames in the *Groundwork*. Still, as she observes, it is 'overwhelmingly plausible to believe' that there is such a rational requirement, and Kant accepts it.³¹

Now Kant's conception of happiness is as 'the sum of satisfaction of all inclinations'³² rather than as the eudaimonist desire to live well, so this principle does not yet count as congruent with TPE. However, the justificatory lacuna here signals that there is a problem with his conception. From the standpoint of eudaimonists, Kant simply has not thought adequately about the rational order needed for a 'sum of satisfaction of desire'. Even if we think individual desires convey normative force to their objects (something Kant does not accept), there is nothing to convey normative force to any principle that we should seek some 'sum of satisfactions' of them. They make conflicting demands on time, energy, and attention, and *normative assessment* is required to adjudicate between those competing demands. The ancients would have thought, and we should agree, that such a system makes *rational* sense only on being reflected upon, shaped, and directed by the end of living well.³³ If so, then Kant's conception of happiness is not stable on reflection, and in reflective form—under the sort of reflective pressures that the ancients believed we are under as practically rational agents—collapses into congruence with TPE.

There is a further Davidsonian point to be made regarding

²⁹ *Groundwork of the Metaphysics of Morals* 4. 415, trans. M. Gregor, in *Groundwork of the Metaphysics of Morals: A German-English Edition*, ed. J. Timmermann, trans. M. Gregor (Cambridge, 1997).

³⁰ Kant's explanation is that it is an imperative by 'natural necessity': it belongs to the 'essence' of 'every human being' (*Groundwork* 4. 415); it represents an indirect duty, he says, because of the temptation to 'transgress duty' if we do not seek our own happiness (4. 399).

³¹ However, she makes no attempt to justify it within the Kantian framework, despite this admission (C. Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford, 2009), § 3.3.5).

³² *Groundwork* 4. 399; see also 4. 405.

³³ Again, this is precisely the point of Socrates' argument against Callicles in *Gorgias*. See M. LeBar, 'Eudaimonist Autonomy', *American Philosophical Quarterly*, 42 (2005), 171–84.

Kant's theory, one relevant to the prospects for finding a counter-example to TPE. Kant is sceptical that we can ascribe even to *ourselves* success in following the moral law as captured in the Categorical Imperative.³⁴ The effects of material or empirical incentives are too pervasive in our practical lives for us to be sure what principle is really governing our choice. *A fortiori* we are never in a position to ascribe with certainty to *others* practical rationality as modelled by the moral law. Instead we must ascribe to them practical rationality as realized in 'dependent' creatures, which, as we have seen, necessarily requires ascribing to them the dominant desire to live well.

In fact Kant's argument in both the *Groundwork* and later (e.g. *Metaphysics of Morals*) supports DPI. The driving insight of that principle is that we can understand others as rational only by deploying the understanding we have of our own rational agency as realized in our own intentional attitudes. Hence there is a commitment to the explanatory primacy of the first-person experience of rational agency in both Davidson's principle and Kant's defence of freedom and the moral law. Though Davidson's argument is driven by considerations drawn from the interpretation of behaviour, Davidson sees much of our 'behaviour' as intentional and as states possible only for creatures subject to the canons of rationality. In a word, he sees much of our behaviour as action rather than mere bodily movement. It is *other beings*, the *speakers* (hence *agents*) about whom Davidson ubiquitously writes, who are the ultimate object of interpretation. Hence, for Davidson and Kant, if we are to see others as practically rational, then we can do so only by using the only model of practical rationality available to us: our own, discovered in first-personal practical agency.³⁵

6. Objections: putative counter-examples

Thus Davidson's constraints on interpretation, coupled with the nature of practical rationality as revealed by the ancients, entail that

³⁴ *Groundwork* 4. 407.

³⁵ A divide on this focus exists among ancient ethical theories as well. The theories on which we focus—those of Plato, Aristotle, and the Stoics—all give a similar primacy to the first-person perspective of agency. Others—those of e.g. Epicurus, the Cyrenaics, and the sceptical schools—do not, which explains in part our setting them aside from our discussion.

we must ascribe the dominant desire to live well to agents. We may seem, however, to have no trouble imagining counter-examples. We conclude by considering six putative counter-examples. Each takes the form of an objection to TPE, which, if successful, invalidates the move to PE.

(a) *The immoral person*

Can we make sense of immoral people, like the violent bank robber who desires to kill the guard so that she can rob the vault, without rejecting TPE? We can. As we have shown, TPE follows from DPI and EPR. For our violent bank robber to ground an objection to TPE, she would have to ground an objection to either DPI or EPR. She does neither.

Consider EPR. While *we* (probably) do not desire to kill the guard so that *we* can rob the bank, if we did, we would best make sense of ourselves as desiring this because we *also* desire to have money, to have things that money can buy, to enjoy the things that money can buy, and many others, all of which bottom out in our desire to live well. We could take *ourselves* to desire to kill the guard only if we understood that desire to be had in the service of the ultimate desire to live well—even though as a matter of actual fact we probably think that this is to misunderstand what living well requires. This is of no consequence. Ancient ethical accounts are replete with cases where the protagonist attempts to show that an interlocutor does not have reason to do what he supposes he would like to do, simply because correct beliefs about how to live well proscribe doing so. The problem is the incorrect beliefs about what living well consists in, not the motives for vicious action.

Might they be a problem for DPI? Were we to enter our neighbourhood bank, we probably would not desire to kill the guard so that we could rob the vault. How can we ascribe immoral desires to the bank robber when (it is to be hoped) we would not ourselves have them? Here Davidson is particularly helpful. We might count robbery as irrational in so far as it violates the law and cheats people out of hard-earned wealth. Davidson's account entails that calling someone 'irrational' makes sense only if we can ascribe to her a background set of beliefs and desires that are themselves rational. Recall his insight that false beliefs can be had only by agents who generally have true ones. False beliefs must be inconsistent with

truths that we hold or otherwise be irrational for us to hold *given* that we take certain other propositions to be true. Davidson explains:

Inconsistency, or other forms of irrationality, can occur only within the space of reasons; inconsistencies are perturbations of rationality, not mere absence of rationality.³⁶

Just as beliefs can be occasionally false only if normally they are true, persons can be occasionally irrational only if normally they are rational. This is another way of putting Davidson's earlier point that a creature that cannot in principle be understood in terms of our own beliefs is a creature without what we mean by thoughts—where, according to Davidson, our beliefs must themselves be largely consistent lest they not be beliefs (for they would lack content).

Now reconsider our violent bank robber. We have to ascribe to her the belief that money is kept in banks, that guards employed by banks if given the chance would try to thwart bank robbers, even that guards once killed *cannot* thwart bank robbers, and so on. These are all beliefs *we* would have and would have to ascribe to the bank robber to make sense of her words and deeds. Likewise we would have to ascribe to her aims such as having money, getting money through violence, and having things that money can buy, on grounds similar to those on which we ascribed justifying ends to Andrea. Even though robbery might itself count as irrational, robbers would overall be anything but, lest we be unable to interpret what they mean or to regard them as possessing rational agency. DPI holds as much for those acting immorally as for others. Since EPR and DPI entail TPE, the case of the immoral person casts no doubt on the truth of TPE.

(b) *The self-destructive person*

We can see in much the same light self-destructive people, including *in extremis* the suicidal. They do not falsify TPE either, since they too fail to falsify EPR or DPI. We can *ourselves* understand how we might be pushed to various forms of self-destruction, even if we in fact are not. As plausible a hypothesis as any is that the motivation to end one's life comes from the (perceived) *frustration*

³⁶ Davidson, *Subjective*, 125.

of the dominant desire to live well. Because one believes that living well in the future is impossible, one is motivated to terminate one's life now.³⁷ We can imagine lesser forms of self-destruction as driven by the dominant desire to live well too. Self-destructive tendencies pose no problem for EPR.

Nor do they for DPI. Even if *we* don't have the desire to harm ourselves, we can ascribe that desire to *others* only if we can fit it into an otherwise rational pattern of beliefs and desires. Only by ascribing to the self-destructive person a background of beliefs and desires that make her come out as rational can we make sense of her as *desiring* to harm herself at all. Only by first imagining that her intentional states approximate what ours would be were we in her shoes, and then readjusting as necessary to make sense of her as rational overall, can we ascribe to her the desire for self-harm. EPR and DPI remain true, and so does TPE.

(c) *The extremely altruistic person*

Stephen Darwall has suggested a different sort of case.³⁸ Consider Sheila, who is motivated to seek the well-being or welfare of others, without seeing her own well-being as reason-giving for her. Darwall's aim is to support the idea that the notion of rational care is conceptually prior to that of welfare or well-being. Thus one could be practically rational, indeed commendably morally (beneficently) motivated, without any desire to live well. If Sheila is intelligible—if we can interpret her as rational—then TPE is in trouble.

Now, as we read him, Darwall isn't claiming that *we ourselves have*

³⁷ Though we are unlikely to call Socrates 'suicidal', this is precisely the principle behind his willingness to accept the verdict of the people of Athens and go willingly to his execution. (See *Crito* 48 c–d, and see further Socrates' claim at *Gorg.* 512 d–e that it is how well one lives, not how long one lives, that matters.) More generally, Stoic doctrine was explicit that suicide can be appropriate when living longer is not conducive to living well. See e.g. Cic. *Fin.* 3. 61; D.L. 7. 130, in A. A. Long and D. N. Sedley, *The Hellenistic Philosophers*, 2 vols. (Cambridge, 1987), i, § 66H; and Stobaeus 111m in B. Inwood and L. Gerson (eds.), *Hellenistic Philosophy: Introductory Readings* (Cambridge, 1997), § 11–95.

³⁸ S. Darwall, *Welfare and Rational Care* (Princeton, 2002), 37 ff. A similar claim is made by R. Kraut, though within the context of a different theory, in *What is Good and Why* (Harvard, 2007), 56 ff. The details of a response on behalf of TPE to Kraut's view will differ somewhat from those relevant to Darwall, but the main concern is the same. For discussion see M. LeBar, 'Development and Reasons: Review of Richard Kraut's *What is Good and Why*', *Philosophical Quarterly*, 58 (2008), 711–19, and *Living Well*, ch. 4. For a Kantian attack on the practical rationality of someone like Sheila, see J. Hampton, 'Selflessness and Loss of Self', *Social Philosophy and Policy*, 10 (1993), 135–65.

Sheila's mindset, only that *someone could*. So he is challenging that our experience generalizes. That is a challenge to DPI. But the idea behind TPE is also at work here in so far as DPI licenses our application of EPR to others. One way to understand Sheila is to see her as thinking a life devoted to the welfare of others *just is* the life most worth living for her. Whether or not we agree, such a notion is intelligible. It is also compatible with TPE, since devoting her life to others is the form that Sheila sees the satisfaction of her desire to live well must take. DPI would therefore license our taking EPR to apply to her.

Consider the options. On the other hand, in so far as we refrain from ascribing to her a dominant desire to live well, we struggle to understand how to see Sheila as rational at all. There either is or is not a point to her commitment to the welfare of others. If there is, and we ask Sheila what that point might be, then Sheila will be capable of rendering her thinking in a form that we can interpret as rational only to the extent that we can embed her aim within a framework of ends which takes its rational structure from the aim of living well. This requires the dominant desire to live well as part of her psychological economy. Seeing Sheila *as rational* forces us to ascribe to her a justificatory structure for her ends in which the dominant desire to live well is essential, just as it is in our own. DPI and with it TPE are vindicated. On the other hand, if there is *not* a point to Sheila's commitment to the welfare of others, then we simply can't interpret her as rational. Sheila turns out not to be an *altruistic person* in any robust sense, for acting altruistically requires acting on reasons.

Can Sheila simply insist, 'Of course what I do has a point. Its point is to increase the well-being of others'? If so, Sheila is committed to seeing the well-being of some people, but not others, as by their nature reason-giving. That sort of arbitrariness renders her unintelligible to us, so this reply is unavailable. Regardless, even if Sheila has *no* reasons (since her behaviour has no point), this does not falsify TPE. For even those who are thus pathological fail to falsify TPE, as we show next.

(d) *The pathological person*

Darwall means Sheila to be an altruistic person, and so to act on reasons. But what should we say about Sheldon—who, like Sheila

on the pathological interpretation, does not desire to live well, but (unlike her) exhibits outright self-destructive tendencies that manifest some deep psychosis? It is not that Sheila (as here imagined) or Sheldon act out of the *wrong* reasons; it is that, properly speaking, they are not acting out of *any* reasons. Both are non-rational in the sense of being beyond the reach of rational discourse. They are both pathological. It just so happens that Sheila's pathology causes her to behave in ways consistent with altruism; Sheldon's, with self-destruction.

This turns out to be just where TPE has its bite. To the extent that we cannot see such human beings as wanting to live well as we do (EPR), which as essential to our own intentional states is required for us to see them as rational (DPI), we cannot see them as rational (TPE). Conversely, TPE claims that if we are to ascribe rationality to someone, we must also ascribe to him or her the dominant desire to live well. But then, by contrapositive reasoning, if we cannot ascribe to someone that desire, then we cannot ascribe rationality to her in the first place. TPE still turns out to be true.

Put differently, TPE is compatible with there being outliers: human beings lacking the dominant desire to live a good life. But it follows from our argument that our relationships with such people cannot be the normal case of engagement with each other as practically rational agents. Such cases drive our practices of institutionalization and other forms of care for those missing essential components of practical rationality. Sheila and Sheldon both require our help but not in the form of rational counsel. EPR, DPI, and TPE are unaffected.

(e) *The akratic person*

Not everyone who fails to be motivated by the dominant desire to live well is pathological, however. Everyone, ourselves included, occasionally suffers from *akrasia* or weakness of the will. But then we would seem not to be required to ascribe the dominant desire to live well to all people, at least not all of the time. Does that falsify TPE?

It does not. Consider the case of ourselves being akratic. Even though we think that we should exercise tonight rather than watch reruns on television, we simply cannot get ourselves off the couch. We are therefore not in this instance living well—and we know it.

Yet for us to be akratic in the first place, we must in the normal course of events act in a way that is not akratic. We cannot have *weak* wills unless we have *wills*—capacities that order our desire-set in such a way as to make us practically rational. And we argued above that the only way in which we could be so rational is if the desire to live well dominates that set. That was the ancient justification for EPR, and *akrasia* does nothing to discount it.

There is a Davidsonian correlate to this point. Davidson rightly claims that inconsistency and other forms of irrationality can occur only within the space of reasons. Our actions can themselves, therefore, be occasionally inconsistent or irrational only if generally they are not. Since being akratic is a way of being irrational, we can be akratic only occasionally if we are to be akratic at all. On this first-person analysis of *akrasia*, the ancients' and Davidson's conclusions converge.

Now if these considerations are true in our own case, then they are certainly true in the case of others. Not only EPR but DPI itself then survives this alleged counter-example. For we can understand others as suffering from weakness of the will only if we can understand them as having wills, which requires our projecting EPR onto them via DPI. Just as we must desire to live well lest we cannot be practically rational, we must take others to desire this also lest we cannot take them to be practically rational. The only way that *akrasia* can be seen as a general phenomenon is if we ascribe intentional states to others that we would have were we in their situations. DPI remains true.

Moreover, if Davidson is right that *we* can be akratic only occasionally if we are to be akratic at all, then we can *ascribe akasia* to *someone else* only occasionally if we are to ascribe *akasia* to her at all. This further supports DPI.

Still, how is *akasia* even possible on our view if the desire to live well is meant to be dominant? Recall that by 'dominant desire' we mean, *inter alia*, a desire that would be called upon to provide sufficient motivation to overwhelm competing motivations. In the akratic's case the call simply goes unheeded. Davidson is right to allow occasional 'perturbations of rationality' to exist within an otherwise rational framework. Our notion of dominant desire must not preclude the possibility of such perturbations, and we were

careful when defining ‘dominant desire’ above to make sure that it does not.³⁹

(f) *Mixed cases*

Some persons, on different occasions or even on the same occasion, exhibit more than one of the traits exemplified by the alleged defeaters of the sorts we have considered. People are complicated, and our account must say something about this. Because there are many possible mixed cases, we shall focus on one in particular. Though the case is fictive, the ancients take it quite seriously; so does Annas. We have in mind Euripides’ Medea.⁴⁰

Medea, daughter of King Aeëtes of Colchis, comes to Jason’s aid when the king commands that Jason retrieve the Golden Fleece. None the less Medea overplays her hand, kills the king’s uncle, and forces Jason and their two children to flee with her to Corinth. Once in Corinth Jason forsakes Medea and her children and marries King Creon’s daughter. Medea in her fury laments: ‘I know that what I am about to do is bad, but anger is master of my plans, which is the source of the greatest troubles for humankind.’⁴¹ She then kills the king and his daughter, as well as—and of particular importance to us—her own children.

Annas observes two ways to think about Medea’s infanticide. Epictetus thinks that anger masters Medea’s plans by pushing her to accept infanticide as the correct act.⁴² He takes Medea to have made an ultimately reasoned (albeit manifestly incorrect) judgement that killing them would be the best way to wreak revenge on her husband and spare her children ridicule. So understood, Medea turns out to be an instance of the case of the immoral person who desires to live well but mistakes what living well consists in. Our reasoning concerning the immoral person then applies here too. Although—it is to be hoped—we would not desire to commit infanticide to wreak revenge and avoid ridicule, if we did, we would

³⁹ Of course, both the ancients and Davidson themselves analyse *akrasia*. On the latter see ‘How is Weakness of the Will Possible?’.

⁴⁰ J. Annas (ed. and comm.), *Voices of Ancient Philosophy: An Introductory Reader* [*Voices*] (Oxford, 2000), 110–16, which prints Eur. *Med.* 1021–80, as well as relevant passages from Epictetus and Galen (see below).

⁴¹ From Annas, *Voices*, 112.

⁴² Epict. *Diss.* 1. 28. 1–9; 2. 17. 17–25, printed in Annas, *Voices*, 114–15. G. Rudebusch, *Socrates* (Malden, Mass., 2009), 137–9, among those who pointed us to Medea, generally agrees with Epictetus’ interpretation.

best make sense of ourselves as desiring this because we ultimately desired to live well. We can make sense of Medea's being driven by revenge and the desire to spare her children ridicule only if we can make sense of Medea's having many other intentional states—states we would have were we in her situation.

So understood, Medea turns out not to be a mixed case at all, and unproblematic for reasons we have already considered. But Annas observes that there is another way to understand Medea. Galen takes Medea to be torn *between* reason and anger, where anger 'dragged her by force forward towards the children—and back again reason pulled her and led her away, and then again anger pulled against this, and then again reason'.⁴³ On Galen's view, reason and anger take turns controlling Medea until anger ultimately wins. Medea's final act cannot be seen as practically rational in any sense, misguided or otherwise. Galen finds Medea's case mixed between what we might consider the normal one of someone who desires to live well and is thereby motivated by that desire, and the pathological one of someone whom it makes no sense to count as practically rational—precisely because it makes no sense to see her, at least in certain instances, as desiring to live well. Medea is, as Galen puts it, 'dragged and pulled', her pathology ultimately victorious. She is therefore in the end non-rational, beyond the reach of rational discourse; she has gone mad with rage. It is not that she acts out of the *wrong* reasons; it is that, ultimately, she is not acting out of *any* reasons.

On this interpretation too there is no problem for TPE. As we have seen, neither the normal nor the pathological case individually falsifies EPR, DPI, or therefore TPE. But can our view account for the possibility of someone who alternates *between* normality and pathology—between, that is, rationality and non-rationality?

The idea of someone's alternating between rationality and non-rationality amounts to the idea of someone being ascribable with the dominant desire to live well at some moments but not others. An example might be Alzheimer's sufferers, whom we describe as having 'good' days and 'bad' days, where, we take it, these are short-hand for days on which we can engage them rationally and those on which we cannot. Correspondingly, we either treat such people as (momentarily) rational, hence as having a dominating desire to live well, or we do not treat them (momentarily) as rational at all. When

⁴³ Galen, *PHP* 3. 3. 13–24, printed in Annas, *Voices*, 113–14.

we do not know whether they are capable of rational self-regulation, we do not know whether argument about what to do is appropriate. Our argumentative resources for engaging them *as persons* with the capacity to respond to reason draw on shared attitudes (DPI), including the dominant desire to live well (EPR). It is simply a fact of life that sometimes we can do this, sometimes not.⁴⁴

Admittedly the case of Medea—and other mixed cases that space prevents our considering—are hard. They should be. When it is unclear to what degree someone is better helped by institutionalization than rational discourse philosophy simply struggles. But so do psychology, literary analysis (why else would Medea receive so much attention among the ancients?), history, and the rest of those disciplines tasked with making sense of the human.

Viewed in this light, Medea turns out to be not so much a counter-example as a reminder that the contours of rationality defy easy definition. But her case does nothing to detract from our argument that considerations of charity in interpretation, and NE in practical rationality, demonstrate that we would be unable to ascribe rationality to anyone without ascribing to her the dominant desire to live well. It would be impossible for us to count as a person anyone to whom we would not ascribe that desire, and possible to count as a person only those to whom we would. We conclude that PE is true: each person has the desire to live well as a dominant element in his or her psychological economy.

Ohio University—Washington and Lee University

BIBLIOGRAPHY

- Annas, J., *The Morality of Happiness* (Oxford, 1993).
 — (ed. and comm.), *Voices of Ancient Philosophy: An Introductory Reader* [*Voices*] (Oxford, 2000).
 Barnes, J. (ed.), *The Complete Works of Aristotle: The Revised Oxford Translation*, 2 vols. (Princeton, 1984).
 Broadie, S., *Ethics with Aristotle* (Oxford, 1991).

⁴⁴ Such cases need not be of organic origin. On ancient views of virtue and vice, vice itself leads to degradation of capacities for rational choice and action in ways that resemble the organic pathologies. Consider Plato's account of the defective constitutions in *Republic* 8 and 9. For discussion of this process in Aristotle's theory see C. Korsgaard, 'Aristotle on Function and Virtue', *History of Philosophy Quarterly*, 3 (1986), 259–79.

- Caldwell, D., 'The Shared-Value Thesis', *Southwest Philosophy Review*, 21 (2006), 61–8.
- Cooper, J. M. (ed.), *Plato: Complete Works* (Indianapolis and Cambridge, 1997).
- Darwall, S., *Welfare and Rational Care* (Princeton, 2002).
- Davidson, D., *Essays on Actions and Events [Essays]*, 2nd edn. (New York and Oxford, 2001).
- 'Expressing Evaluations', in *Problems*, 19–38.
- 'How is Weakness of the Will Possible?', in *Essays*, 21–42.
- *Inquiries into Truth and Interpretation [Inquiries]* (Oxford, 2001).
- 'Intellectual Autobiography of Donald Davidson', in Hahn (ed.), *Davidson*, 3–70.
- 'The Interpersonal Comparison of Values', in *Problems*, 59–74.
- 'Mental Events', in *Essays*, 207–28.
- 'Moods and Performance', in *Inquiries*, 109–22.
- 'On the Very Idea of a Conceptual Scheme', in *Inquiries*, 183–98.
- *Problems of Rationality [Problems]* (Oxford, 2004).
- *Subjective, Intersubjective, Objective [Subjective]* (Oxford, 2002).
- 'Thought and Talk', in *Inquiries*, 155–70.
- *Truth and Predication* (Harvard, 2005).
- *Truth, Language, and History* (Oxford, 2005).
- Fodor, J. A., *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass., 1989).
- Foot, P., *Natural Goodness* (Oxford, 2001).
- Goldberg, N., 'E pluribus unum: Arguments against Conceptual Schemes and Empirical Content', *Southern Journal of Philosophy*, 42 (2004), 411–38.
- 'Swampman, Response-Dependence, and Meaning', in G. Preyer (ed.), *Davidson's Philosophy: Truth, Meaning, and the Mental* (Oxford, forthcoming).
- 'Tension within Triangulation', *Southern Journal of Philosophy*, 46 (2008), 367–83.
- 'Triangulation, Untranslatability, and Reconciliation', *Philosophia*, 37 (2009), 261–80.
- Hahn, L. E. (ed.), *The Philosophy of Donald Davidson [Davidson]* (La Salle, Ill., 1999).
- Hampton, J., 'Selflessness and Loss of Self', *Social Philosophy and Policy*, 10 (1993), 135–65.
- Hard, R. (trans.), and Gill, C. (ed.), *The Discourses of Epictetus. The Handbook. Fragments* (London, 1995).
- Hursthouse, R., *On Virtue Ethics* (Oxford, 1999).
- Inwood, B., and Gerson, L. (eds.), *Hellenistic Philosophy: Introductory Readings* (Cambridge, 1997).

- Kamtekar, R., 'Plato on the Attribution of Conative Attitudes', *Archiv für Geschichte der Philosophie*, 88 (2006), 127–62.
- Kant, I., *Groundwork of the Metaphysics of Morals: A German–English Edition*, ed. J. Timmermann, trans. M. Gregor (Cambridge, 1997).
- *Religion within the Boundaries of Mere Reason*, trans. A. W. Wood and G. Di Giovanni, in Kant, *Religion and Rational Theology* (Cambridge, 1996).
- Korsgaard, C., 'Aristotle on Function and Virtue', *History of Philosophy Quarterly*, 3 (1986), 259–79.
- *Self-Constitution: Agency, Identity, and Integrity* (Oxford, 2009).
- Kraut, R., *What is Good and Why* (Harvard, 2007).
- LeBar, M., 'Development and Reasons: Review of Richard Kraut's *What is Good and Why*', *Philosophical Quarterly*, 58 (2008), 711–19.
- 'Eudaimonist Autonomy', *American Philosophical Quarterly*, 42 (2005), 171–84.
- *The Value of Living Well [Living Well]* (Oxford, forthcoming).
- Long, A. A., and Sedley, D. N., *The Hellenistic Philosophers*, 2 vols. (Cambridge, 1987).
- Martin, B., 'Interpretation and Responsibility: Excavating Davidson's Ethical Theory', in Hahn (ed.), *Davidson*, 345–56.
- 'The Moral Atmosphere: Language and Value in Davidson', *Southwest Philosophy Review*, 6 (1990), 89–97.
- Rovane, C., 'On the Very Idea of an Ethical Scheme', *Iyyun*, 54 (2005), 109–28.
- Rudebusch, G., *Socrates* (Malden, Mass., 2009).